CLASSIFICATION OF HEPATIC LESIONS FROM CT IMAGES USING TEXTURE FEATURES AND NEURAL NETWORKS

M. Gletsos[1], S. G. Mougiakakou[1*], G. K. Matsopoulos[1], K. S. Nikita[1], A. S. Nikita[2], D. Kelekis[2]
[1]Department of Electrical and Computer Engineering, National Technical University of Athens, Greece
[2]Second Department of Radiology, University of Athens, Medical School, Greece
*e-mail: smougia@cc.ece.ntua.gr

Abstract-**In this paper a computer-aided diagnostic system for the classification of hepatic lesions from Computed Tomography (CT) images is presented. Regions of Interest (ROI's) taken from non-enhanced CT images of normal liver, hepatic cysts, hemangiomas, and hepatocellular carcinomas (a total of 147 samples), have been used as input to the system. The system consists of two levels: the feature extraction and the classification levels. The feature extraction level calculates the average grey scale and 48 texture characteristics, which are derived from the spatial grey-level co-occurrence matrices, obtained from the ROI's. The classifier level consists of three sequentially placed feed-forward Neural Networks (NN's), which are activated sequentially. The first NN classifies into normal or pathological liver regions. The pathological liver regions are classified by the second NN into cysts or "other disease". The third NN classifies "other disease" into hemangiomas and hepatocellular carcinomas. In order to enhance the performance of the classifier and improve the execution time, the dimensionality of the initial feature vector has been reduced using the sequential forward floating selection method for each individual NN input vector. A total classification rate of 98% has been achieved.**

***Keywords*–Texture features, sequential floating forward selection, neural network, classification, liver CT images**

I. INTRODUCTION

The use of medical imaging and tissue characterisation techniques is becoming increasingly popular in diagnosis, treatment and research. However, in the case of liver disease, characterisation of hepatic lesions from Computed Tomography (CT) or B-Scan Ultrasound is troublesome, even for experienced radiologists, and often requires confirmation by biopsies, or other invasive techniques [1]. To overcome this, a variety of computer-aided diagnostic systems have been proposed. The system involves use of different measures, such as texture features, grey scale, fractal dimension estimators or shape descriptors, combined with a classifier [2-4]. Recently, it has been reported that texture features combined with Neural Networks (NN's) can improve the performance of the classification of hepatic tissues. More specifically, in [3], a classification rate of about 83% has been achieved using texture features with a probabilistic NN classifier, in order to classify hepatoma and hemangioma regions taken from CT images, while in [4] the accuracy was 100% using texture features combined with a 2 layered feed-forward NN classifier for the classification of ultrasound images into normal tissue, hemangioma and malignancy.

In this paper, a diagnostic system has been developed, in order to classify four (4) hepatic tissue types: normal liver (C1), hepatic cyst (C2), hemangioma (C3), and hepatocellular carcinoma (C4).

II. METHODOLOGY

The proposed diagnostic system is presented in Fig. 1. It consists of two levels: the feature extraction level and the classifier levels.

A. Data

The Second Department of Radiology of the Medical School of Athens, Greece has provided the data used for the training of the diagnostic system. The acquired non-enhanced CT images have a resolution of 512×512 pixels, with 8-bit grey-level at the W150+60 window, and they have been taken from both patients and healthy controls. All the hepatic lesions have been validated by needle biopsies, density measurements and the typical pattern of enhancement after the intravenous injection of iodine contrast. Experienced radiologists have delineated the position, size and extent of the lesions. A total of 147 free-hand regions of interests (ROI's) have been sampled, 76 of which belong to healthy controls, 19 to cysts, 28 to hemangioma and 24 to hepatocellular carcinoma.

B. Feature Extraction

For each ROI, the feature extraction sub-system calculates a 49-dimensional feature vector consisting of the average grey scale of the ROI, and 48 texture characteristics. The texture characteristics are derived from eight texture features, calculated from the spatial grey-level co-occurrence matrices of the ROI's [5]. These features correspond angular second moment (*asm*), contrast (*con*), correlation (*corr*), sum of squares (*ss*), inverse difference moment (*idm*), entropy (*ent*), homogeneity (*hg*), cluster

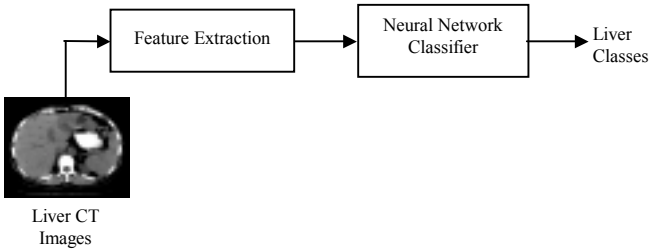


Fig. 1. Description of the proposed system of hepatic lesions from CT images.

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle Classification of Hepatic Lesions From CT Images Using Texture Features and Neural Networks	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Department of Electrical and Computer Engineering National Technical University of Athens Greece	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 803 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.

Abstract

Subject Terms

Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU

Number of Pages
4

tendency (*clt)* and are dependent on pixel spacing and angular direction. They are calculated for six different values of pixel spacing (d), ranging from 1 to 12 pixels. For each value of pixel spacing, the feature values are computed by averaging over four uniformly distributed angular directions, 0^o, 45^o, 90^o, and 135^o. Thus, a total of 48 texture characteristics are obtained.

C. Feature Selection

Due to the high dimensionality of the feature vectors, dimensionality reduction has been performed with use of feature selection. Feature selection methods select a subset of the original set of features. They involve use of search algorithms and the maximisation of a given evaluation function, which corresponds to the separability between the class populations. In this work, the evaluation function J_1 used by the algorithms is a distance criterion, based on the scatter matrices:

$$J_1 = \text{trace}\,\{S_w^{-1} S_b\} \quad (1)$$

where S_w and S_b are respectively the within-class and between-class scatter matrices, calculated from the covariance matrices of the class populations, as defined in [6].

A variety of search algorithms have been tested for feature selection. The initial tests have shown that optimal search methods, such as exhaustive search, and branch-and-bound algorithms [6], were computationally prohibitive, due to the dimensionality of the feature space. Consequently, heuristic algorithms have been applied. Heuristic algorithms search partly the data space, and they provide results close to the optimal at low computational cost [7]. The most popular categories of heuristic algorithms are forward selection and backward selection.

The backward or "top-down" methods, including sequential backward selection (SBS) [7], start with the full feature set and gradually eliminate features. Due to the high dimensionality of the feature subsets at the first stages, they were susceptible of producing unreliable results during the inversion of large matrices.

Forward or "bottom-up" methods, including sequential forward selection (SFS) and sequential forward floating selection (SFFS) [8], start with an empty feature subset and sequentially add features. SFS, which is one of the most commonly used methods, adds one feature at each stage, which when combined with the previously selected features gives the best measure of the evaluation function. On the other hand, the SFFS method is allowed to "float", with use of conditional SFS and SBS at each stage.

The SFFS method has been finally applied for the reduction of feature dimensionality, because contrary to SFS, it can detect "nested" features that otherwise remain hidden, and provide a solution closer to the optimal and at affordable computational cost [7, 8]. The steps of the SFFS algorithm can be summarised as follows:

Suppose k features, x_1, x_2,...x_k, have already been selected, to form a feature set X_k.

Step 1:Inclusion. With plain SFS add one more feature x_{k+1} to the set X_k of selected features, to form $X_{k+1}=\{X_k+x_{k+1}\}$.

Step 2: Conditional exclusion. With plain SBS select one feature x_r from the set X_{k+1}. If x_r is the previously selected feature ($x_r=x_{k+1}$), then $k=k+1$ and return to Step 1; else exclude it from the set of features and form a temporary set $X_k'=\{X_{k+1}-x_r\}$ and go to Step 3, unless $k=2$. If $k=2$ set $X_k=X_k'$ and return to Step 1.

Step 3: Continuation of conditional exclusion. With plain SBS select one feature x_s from the temporary set X_k'. If the resulting subset $\{X_k'-x_s\}$ is not the best performing subset in terms of the evaluation function J_1, then set $X_k=X_k'$ and return to Step 1; else let $X_{k-1}'=\{X_k'-x_s\}$, $k=k-1$, and repeat Step 3. If $k=2$ set $X_k=X_k'$ and return to Step 1.

The algorithm is initialised with $k=0$ and use of SFS until $k=2$.

D. Neural Network Classifier

The second level of the diagnostic system is a classifier implemented with NN techniques. NN's have been applied successfully in classification problems in the field of medical imaging. In the current application, the proposed classifier consists of three sequentially placed feed-forward NN's, as shown in Fig. 2. More specifically, the first NN (NN1) classifies to normal and pathological liver regions. If the region is found to be pathological, the second NN (NN2) is activated in order to classify to regions with cyst or other disease. If the NN2 diagnosis is cyst, the process is terminated, else the third NN (NN3) is activated in order to classify between hemangioma and hepatocellular carcinoma. This architecture has been selected for two main reasons: Firstly, it is very close to the architecture of the biological systems, where sequentially placed layers of physiological neurons are organised in a hierarchical way, in order to solve stepwise more complex problems. Secondly, from our experiments it has been found that the proposed architecture is more suitable to the specific problem [9].

The back-propagation algorithm with adaptive learning rate and momentum has been used in order to train the three NN's [10, 11]. Each NN consists of one input layer with number of neurons equal to the number of the selected features, one hidden layer with number of neurons varying from 2 to 20, and one output layer of one neuron for the classification of the liver region. The output neuron of each NN has two possible states, 0 and 1, according to the diagnosis of the hepatic tissue.

The initial weights of all NN's have been randomly selected in the range [–1.0, +1.0]. Because the log-sigmoid activation function has been used for the hidden layer, the data of input vectors of each NN have been normalised in the range [–1.0, +1.0] before their presentation to the corresponding NN's. The tan-sigmoid activation function has been used for the output layer of each NN's.

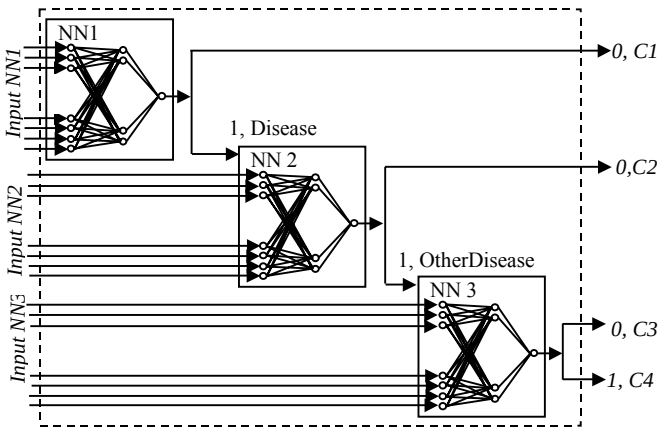


Fig. 2. The architecture of the implemented NN classifier.

The appropriate number of hidden neurons and the values of the learning rate and the momentum have been estimated using a process of trial-and-error until no further improvement in classification could be obtained.

In order to avoid overtraining and achieve an acceptable generalisation of the classification, three data sets have been selected for each NN's: training sets, validation sets, and testing sets. The NN's are trained using the training set and the training phase stops when the performance in the validation set is maximised. The generalisation ability is tested using the testing set, which contains samples that have not been used previously.

III. Results

From a total of 147 samples, the full set has been used for training and testing of NN1, whereas 71 and 52 samples have been used for training and testing of NN2 and NN3, respectively. The training, validation and testing sets for each NN have been selected as shown in Table I.

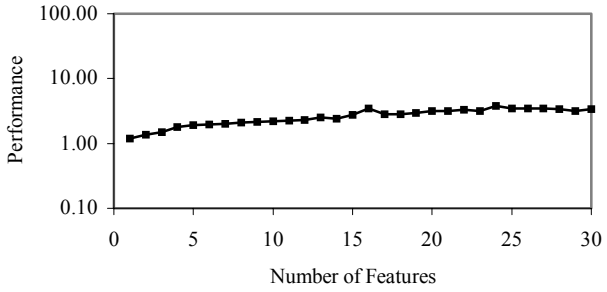
After the application of the SFFS method, as described in Section II, the resulted input vectors have been fed to the NN's comprising the classifier of the system. The final design of the proposed system, according to the maximisation of the classification performance, is given in the following paragraphs.

Neural Network 1: As shown in Fig. 3a, the first apparent "plateau" of the evaluation function J_1 corresponds to subsets of 8 or 9 features. Although subsets with more features might show higher performance, the choice of feature subsets has been biased in favour of small input vectors. Thus, the NN has been trained with two different subsets containing 8 and 9 features respectively. Classification rate of 100%, as shown in Table II, has been achieved using an input vector with 8 features and a hidden layer with 4 neurons, whereas the performance using the full 49-feature vector was 97%.

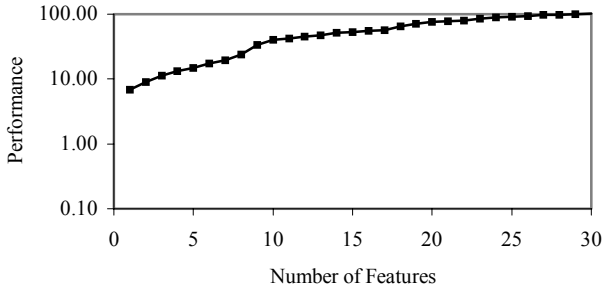
Neural Network 2: The second NN has been trained with 3 datasets containing different number of features. Subsets with 3, 4, and 10 features correspond to "plateaus" in the diagram shown in Fig. 3b. All selected feature subsets produced 100% correct classification rate, not only in the training, but also in the validation and testing sets. This behaviour has been expected, since discriminating cyst from

TABLE I
DIVISION OF THE TOTAL DATABASE IN TRAINING, VALIDATION, AND TESTING SETS

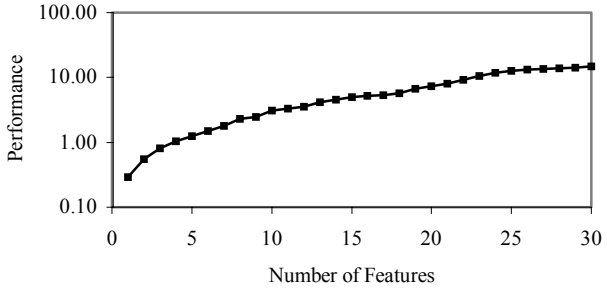
		Training Set	Validation Set	Testing Set
NN1	Normal	43	17	17
	Disease	41	15	15
NN2	Cyst	11	4	4
	Other Disease	30	11	11
NN3	Hemangioma	16	6	6
	Hepatocellular carcinoma	14	5	5



(a)



(b)



(c)

Fig. 3. Evaluation function J_1 versus subset dimensionality during feature selection with SFFS for the classes (a) normal –diseased, (b) cyst – other diseases and (c) hemangioma – hepatocellular carcinoma.

TABLE II
CLASSIFICATION PERFORMANCE OF THE NN'S ACCORDING TO THE SIZE OF THE FEATURE VECTOR

NN1			*NN2*			*NN3*		
Number of features	*Validation Set*	*Testing Set*	*Number of features*	*Validation Set*	*Testing Set*	*Number of features*	*Validation Set*	*Testing Set*
49	91%	97%	49	100%	100%	49	91%	82%
9	100%	100%	10	100%	100%	10	82%	91%
8	100%	100%	4	100%	100%	5	82%	91%
			3	100%	100%	4	82%	91%
						3	82%	73%

"other" lesion types is fairly easy, and the evaluation function J_1 has accordingly reached high values. The same classification performance has been achieved using the initial feature vector. The requirement to achieve the highest classification rate with the smallest feature vector has resulted in the selection of 3 input neurons, implemented with 4 hidden neurons, as shown in Table II.

Neural Network 3: Input vectors consisting of 3, 4, 5 or 10 features, according to Fig. 3c, have been used for the third NN. After the training, validation and testing with those datasets it has been found that the performance of the NN3 is highest (91%) using 4 input neurons (Table II) and 11 hidden neurons, while the performance of NN3 with the 49 dimensional feature vector was 82%.

Total System Performance: The requirement for the highest classification rate with the smallest feature vectors is satisfied with the selection of 8 features for NN1 (gray scale, *asm*, *corr*, *ss*, and *idm* for d=1, *ss* for d=6, *corr* for d=8, and *clt* for d=12), 3 features for NN2 (gray scale, *hg* for d=2 and d=12), and 4 for NN3 (*corr* for d=2, and d=12, *ss* for d=4 and *ent* for d=6). The integration of the final system, which consists of the distinct NN's achieving best classification of the hepatic lesions, has resulted in high classification rates. Analytically, a total classification rate of 100% in the training set, 97% in the validation set, and 98% in the testing set have been achieved. Table III shows the confusion matrix of the classifier.

IV. Conclusion

A computer-aided diagnostic system has been proposed for the classification of hepatic lesions from CT images. The use of texture features, the application of dimensionality reduction with sequential forward floating selection (SFFS), and the choice of a classifier consisting of three sequentially placed neural networks, have resulted in a total classification performance of 98%, for four classes of hepatic tissue.

TABLE III
CONFUSION MATRIX OF THE NN CLASSIFIER

	Real Diagnosis							
	Validation Set				*Testing Set*			
NN Diagnosis	*C1*	*C2*	*C3*	*C4*	*C1*	*C2*	*C3*	*C4*
C1	17	0	0	0	17	0	0	0
C2	0	4	0	0	0	4	0	0
C3	0	0	4	0	0	0	5	0
C4	0	0	2	5	0	0	1	5

The application of feature selection with SFFS in the implemented system has resulted not only in high classification performance, but also in reduced complexity, and execution times in the order of a few seconds. Use of larger databases of CT images is expected to improve the system robustness and ensure the repeatability of the resulted performance. The proposed system can be extended for other types of images or other classes of liver disease, provided that the feature vectors are re-evaluated and the NN's are re-trained.

Acknowledgement

The authors express their gratitude to the General Secretariat for Research and Technology of the Ministry of Development for supporting the research presented in this paper under the PENED99 project.

References

[1] H.M. Taylor, P.R. Ros, "Hepatic Imaging: An Overview," *Radiologic Clinics of North America*, vol.36, no. 2, March 1998.
[2] Y.N. Sun, M.H. Horng, X.Z. Lin, and J.Y. Wang, "Ultrasonic Image Analysis for Live Diagnosis," *IEEE Engineering in Medicine and Biology*, vol. 11-12, pp. 93-101, 1996.
[3] H. Sujana, S. Swarnamani, and S. Suresh, "Application of Artificial Neural Networks for the Classification of Liver lesions by Image Texture Parameters," *Ultrasound in Medicine and Biology*, vol. 22, no. 9, pp. 1177-1181, 1996.
[4] E.L. Chen, P-C. Chung, C.L. Chen, H.M. Tsa, and C.I. Chang, "An Automatic Diagnostic System for CT Liver Image Classification," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 6, pp. 783-794, 1998.
[5] R.M. Haralick and L.G. Shaphiro, *Computer and Robot Vision*, vol. I. Addison-Wesley Publishing Company, 1992.
[6] K. Fukunanga, *Introduction to Statistical Pattern Recognition*, 2nd ed., Academic Press Inc., 1990.
[7] A.K. Jain, R.P.W. Duin, and J. Mao, "Statistical Pattern Recognition: A Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-37, January 2000.
[8] P. Pudil, J. Novovičová, and J. Kittler, "Floating Search Methods in Feature Selection," *Pattern Recognition Letters*, vol. 15, pp. 1119-1125, November 1994.
[9] S.G. Mougiakakou, M. Gletsos, G.K. Matsopoulos, A.S. Nikita, K.S. Nikita, and D. Kelekis, "Development of a Decision Support System for the Classification of Hepatic Lesions from CT Images," *4th Inter. Conf. Neural Networks and Expert Systems in Medicine and Healthcare*, June 2001 (in press).
[10] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice-Hall, 1999.
[11] T. Masters, *Practical Neural Network Recipes in C++*, Academic Press, 1993.